

***FileNet*: A WEB-BASED RESEARCH REPOSITORY
SYSTEM FOR IS-FIT UUM**

ZAHRAH BT YAHYA

UNIVERSITI UTARA MALAYSIA (2006)

***FileNet*: A WEB-BASED RESEARCH REPOSITORY
SYSTEM FOR IS-FIT UUM**

A thesis submitted to the Graduate school in partial
fulfillment of the requirement for the degree Master of Science (Information Technology)
Universiti Utara Malaysia

By

Zahrah bt Hj. Yahya

© Zahrah Yahya, September 2006. All right reserved



PUSAT PENGAJIAN SISWAZAH
(Centre For Graduate Studies)
Universiti Utara Malaysia

PERAKUAN KERJA KERTAS PROJEK
(Certificate of Project Paper)

Saya, yang bertandatangan, memperakukan bahawa
(I, the undersigned, certify that)

ZAHHRAH BT YAHYA

calon untuk Ijazah
(candidate for the degree of) **MSc. (Information Technology)**

telah mengemukakan kertas projek yang bertajuk
(has presented his/ her project paper of the following title)


FILENET: A WEB-BASED RESEARCH REPOSITORY SYSTEM

seperti yang tercatat di muka surat tajuk dan kulit kertas projek
(as it appears on the title page and front cover of project paper)

bahawa kertas projek tersebut boleh diterima dari segi bentuk serta kandungan dan meliputi bidang ilmu dengan memuaskan.
(that the project paper acceptable in form and content, and that a satisfactory knowledge of the field is covered by the project paper).

Nama Penyelia Utama
(Name of Main Supervisor): **Associate Professor Dr. Wan Rozaini Sheik Osman**

Tandatangan
(Signature)

: 

Tarikh
(Date)

: 4/12/06

PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for the postgraduate degree from Universiti Utara Malaysia, I agree that University Library may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purpose may be granted by my supervisor or, in their absence by the Dean of the Graduate School. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to Universiti Utara Malaysia for any scholarly use which may be made of any material from my thesis.

Request for permission to copy or to make other use of materials in this thesis, in whole or in part, should be addressed to:

**Dean of Faculty of Information Technology
Universiti Utara Malaysia
06010 UUM Sintok
Kedah Darul Aman**

ABSTRAK

Sistem Capaian Maklumat berasaskan aplikasi laman web telah banyak dilaksanakan di dalam organisasi khususnya di dalam institusi pendidikan yang bertaraf universiti. Objektif *fileNet* dibangunkan adalah untuk menyediakan laman web yang efektif bagi kegunaan para penyelidik untuk mencari, mencapai dan menggunakan maklumat yang ada di dalam laman web tersebut, sekaligus dapat membantu penyediaan kajian yang lebih produktif dan berkualiti. *FileNet* telah dibangunkan untuk Jabatan Sistem Maklumat di Fakulti Teknologi Maklumat Universiti Utara Malaysia dan dibangunkan dengan menggunakan perisian PHP, MySQL dan Apache. Unified Modeling Language (UML) telah digunakan bagi mereka bentuk model keperluan sistem. Kertas Kajian ini juga turut membincangkan metodologi 'web system' yang dicadangkan oleh Ginige. Kajian ini telah merumuskan bahawa *fileNet* merupakan sistem yang efisien bagi simpanan dan capaian fail dan juga dapat meningkatkan kerjasama di antara kumpulan kerja melalui fungsi dan organisasi.

ABSTRACT

Information Retrieval system using web-based application has been implemented in many organization especially education institution at university level. The objectives of *fileNet* is to build effective research tools for researchers to access, explore and use the information, which may lead to improve research productivity and quality. *fileNet* has been developed for Information System department Universiti Utara Malaysia by using PHP, MySQL and Apache. The requirement model for *fileNet* was designed using the Unified Modeling Language (UML). This paper also discussed the web system methodology as proposed by Ginige as the methodology for system development. This paper concludes *fileNet* is the reliable and efficient for file storage and retrieval and also can enhance workgroup collaboration across functions and organization

ACKNOWLEDGEMENT

First of all I am so thankful to Allah for giving me the courage, ability, and strength to complete this project.

I would like to express my gratitude sincere appreciation to both of my supervisor Associate Professor Dr. Wan Rozaini Sheik Osman for guiding the research presented in this dissertation. They have been a constant source of motivation and encouragement. I appreciate and thank them for their continuous support, for being always accessible and for providing invaluable feedback on my work. Their encouragement helped shape the direction of my work. Without them, maybe this project is not completed properly.

I give my warm thanks to my friends for their support, encouragement, and collaboration in sharing ideas during this project. The persons that I mentioned above are my husband Mohd Nizam Bin Husen and my friends Wan Hazimah bt Wan Ismail and Muhammad bin Maharoff. Thanks for being a wonderful person to me.

I would also like to thank to my beloved family especially my beloved parents, Hj. Yahya bin Hj. Ibrahim and Hjh Umaimah bt Hj. Kasim who never stop giving me a support, love, prayers, and always been there for me. Your love gives me strength to complete this study. Thanks to all my sisters and brothers for your continuous support.

Finally, I thank to everybody that involved in my project direct or indirectly for productive interactions and for their support.

TABLE OF CONTENTS

PERMISSION TO USE	i
ABSTRAK	ii
ABSTRACT	iii
ACKNOWLEDGEMENT	iv
TABLE OF CONTENTS	v
LIST OF FIGURES	viii
LIST OF TABLES	x
LIST OF ABBREVIATIONS	xii

CHAPTER 1: INTRODUCTION

1.1 Overview of Study	1
1.2 Problem Statement	3
1.3 Objectives of Study	4
1.4 Scope of Study	5
1.5 Significance of Study	6
1.6 Organization of the Report	6

CHAPTER 2: LITERATURE REVIEW

2.1 World Wide Web	8
2.2 Web-based Application	10
2.3 Architecture of Web Application	10
2.4 The Advantage of Web-based Application	11
2.5 Web Development Tools	12
2.5.1 MySQL Database	13
2.5.2 PHP Scripting Language	14
2.5.3 Apache Web Server	14
2.6 Web-Based Application for Document Search and Retrieval Adoption	15

CHAPTER 3: RESEARCH METHODOLOGY

3.1 Introduction	22
3.2 Context Analysis	22
3.3 Product Model	23
3.3.1 System Architecture	23
3.3.2 Application Architecture	24
3.3.3 Software Architecture	25
3.4 Website Development	26
3.4.1 Web Site Design	27
3.4.2 Web Site Construction	28
3.5 Testing	29

CHAPTER 4: FINDINGS

4.1 Requirement Model	31
4.1.1 Rational Rose 2000	32
4.1.2 Unified Modeling Language (UML)	32
4.1.3 The UML Diagram	33
A. Business Use Case Model	34
B. Use Case Diagram	35
C. Activity Diagram	39
D. Sequence Diagram	44
E. Collaboration Diagram	53
F. Class Diagram	61
4.2 Prototype Development	63
4.2.1. The Development Tools	63
4.2.1.1 Macromedia Dreamweaver MX	63
4.2.1.2 PHP:Hypertext Processor	66
4.2.1.3 MySQL	68
4.2.2 Preview of the Prototype	70
4.3 System Testing	76
4.3.1 System Requirement Testing and Results	76
4.4 Analysis of Results	79

CHAPTER 5: CONCLUSION & RECOMMENDATION

5.1 Conclusion	80
5.2 Limitation and Recommendation	81

REFERENCES

REFERENCES	83
-------------------	-----------

APPENDIXES

APPENDIX A	87
-------------------	-----------

LIST OF FIGURES

Figure 2.1	Architecture of Web application	
	(Source: Greenspan & Bulger, 2001)	11
Figure 2.2	Replx EDM Enterprise Document Management interface	16
Figure 2.3	CiteSeer Scientific Literature Digital Library interface	17
Figure 2.4	Current Content: Journal Search	18
Figure 2.5	DBLP: The Collection of Computer Science Bibliographies	20
Figure 3.1	Development Process for Web-based Application	
	proposed by Ginige	21
Figure 3.2	Application architecture of the <i>fileNet</i>: Web-based Research	

	Repository System for IS-FIT UUM	24
Figure 3.3	Software architecture	25
Figure 3.4	Two-stage approaches to Design and Construction of Web Sites	27
Figure 3.5	The sample design of <i>fileNet</i>	28
Figure 4.1	The use case diagram for <i>fileNet</i>	34
Figure 4.2	Activity Diagram for <i>upload file</i> use case	40
Figure 4.3	Activity Diagram for <i>login</i> use case	41
Figure 4.4	Activity Diagram for <i>search</i> use case	42
Figure 4.5	Activity Diagram for <i>download</i> use case	43
Figure 4.6	Sequence Diagram for <i>upload file</i> (Best Case Scenario)	45
Figure 4.7	Sequence Diagram for <i>upload file</i> (Failure Case Scenario)	46
Figure 4.8	Sequence Diagram for <i>login</i> (Best Case Scenario)	47
Figure 4.9	Sequence Diagram for <i>login</i> (Failure Case Scenario)	48
Figure 4.10	Sequence Diagram for <i>search</i> (Best Case Scenario)	49
Figure 4.11	Sequence Diagram for <i>search</i> (Failure Case Scenario)	50
Figure 4.12	Sequence Diagram for <i>download</i> (Best Case Scenario)	51
Figure 4.13	Sequence Diagram for <i>download</i> (Failure Case Scenario)	52
Figure 4.14	Collaboration Diagram for <i>upload</i> (Best Case Scenario)	53
Figure 4.15	Collaboration Diagram for <i>upload</i> (Failure Case Scenario)	54
Figure 4.16	Collaboration Diagram for <i>login</i> (Best Case Scenario)	55
Figure 4.17	Collaboration Diagram for <i>login</i> (Failure Case Scenario)	56
Figure 4.18	Collaboration Diagram for <i>search</i> (Best Case Scenario)	57

Figure 4.19	Collaboration Diagram for <i>search</i> (Failure Case Scenario)	58
Figure 4.20	Collaboration Diagram for <i>download</i> (Best Case Scenario)	59
Figure 4.21	Collaboration Diagram for <i>download</i> (Failure Case Scenario)	60
Figure 4.22	Class Diagram	61
Figure 4.23	Macromedia Dreamweaver 8	64
Figure 4.24	The structure of <i>fileNet</i> web page using Macromedia Dreamweaver MX	65
Figure 4.25	PHP: Hypertext Processor	66
Figure 4.26	<i>uploads</i> table structure	67
Figure 4.27	Welcome screen	70
Figure 4.28	Login Page	71
Figure 4.29	Upload Page	72
Figure 4.30	Search Page	73
Figure 4.31	Search Result and Download Page	74

LIST OF TABLES

Table 3.1	System Testing Requirement	30
Table 4.1	<i>fileNet</i> modules	67
Table 4.2	Actual System Performance	77

LIST OF ABBREVIATIONS

ASP	Active Server Pages
CSS	Cascading Style Sheets
FIT	Faculty of Information Technology
HTML	Hypertext Markup Language
HTTP	Hypertext Transfer Protocol
IIS	Internet Information Services
IS	Information System
IT	Information Technology
OS	Operating System
PHP	PHP: Hypertext Pre Processor
RDBMS	Relational Database Management System
SQL	Structured Query Language
UML	Unified Modeling Language
URL	Uniform Resource Locator
UUM	Universiti Utara Malaysia
WWW	World Wide Web

CHAPTER ONE

INTRODUCTION

This chapter presents the main idea of this study which is web-based research repository system that is useful for document store, search and retrieval on the web. In addition this chapter also discusses about problem statement, objectives of study, scope of study and the significance of study.

1.1 OVERVIEW OF THE STUDY

In recent years organizations have made huge, conscious, increases in Web-based application investments, reaching into almost every aspect of organizational work. The effect of the Web on many firms has been dramatic (Barry, 2002). Today organization are becoming increasingly aware of the need to properly manage document-based enterprise content that has a direct bearing on workforce productivity and business efficiency. After all, 75 percent or more of enterprise document is unstructured and unmanaged. These documents come in various forms such as office documents, scanned images, application reports, email and drawings (Kobayashi and Takeda, 2000).

In a preliminary interview with Associate Professor Dr Wan Rozaini, the head of Information System (IS) department, Faculty of Information Technology (FIT) in Universiti Utara Malaysia (UUM) has mentioned that all the research documents

The contents of
the thesis is for
internal user
only

REFERENCES

- Antelman, L., & Mendelzon, A. (1998). *Database Techniques for the World Wide Web: A Survey*. ACM SIGMOD record.
- Apache HTTP server version 1.3 (2003). *Documentation*. Available FTP:
<http://httpd.apache.org/docs/>.. Retrieved August 21, 2006
- Barry, C. (2002). *Web-Based Information System- Time for the Revisionist*. Retrieved June 6, 2006, from ACM database.
- Bennet, S., McRobb, S., & Farmer, R. (2002). Object oriented system analysis and design using UML (2nd ed.) McGraw-Hill Education: Backshire.
- Bentley, R., Appelt, W., Busbach, U., Hinrichs, E., Kerr, D., Sikkell, K., & Trevor, G. (1997). Basic Support for Cooperative Work on the World Wide Web. *International Journal of Human Computer Studies: Special issue on Novel Applications of the WWW*. Retrieved June 6, 2006, from ACM database.
- Booch, G., Jacobson, I., & Rumbaugh, J. (2001). *The unified modeling language user guide*. Addison-Wesley: Boston.
- Chidlovskii, B., Roncancio, C., & Schneider, M. (1999). *Semantic Cache Mechanism for Heterogeneous Web Querying*. Retrieved June 6, 2006, from ACM database.
- Computer and Information Science Papers Citeseer Publications ResearchIndex (2005). *CiteSeer. IST Scientific Literature Digital Library*. Retrieved June 7, 2006, from <http://citeseer.ist.psu.edu/citeseer>.

Conallen, J. (2000). Building Web Applications with UML. The Addison-Wesley Object Technology Series.

Deri, L. (1999). Desktop versus Web-based Network Management. *International Journal of Network Management*. Retrieved June 6, 2006, from ACM database.

Elmasri, R., Navathe, S. (2000). Fundamental of Database Systems. (3rd edition). Addison-Wesley.

Ginige, A. (2002). *Web Engineering: Managing the Complexity of Web Systems Development*. Retrieved April 11, 2006, from ACM database.

Ginsburg, M., & Kambil, A. (1999). Annotate: A Web-based Knowledge Management Support System for Document Collections. *Proceedings of the 32nd Hawaii International Conference on System Science-1999*. Retrieved April 11, 2006, from ACM database.

Greenspan, J., & Bulger, B. (2001). *MySQL / PHP Database Applications*. USA: M & T.

Hedja, P. (2000) Architectural Model for User Interfaces of Web-based Applications, 6th *ERCIM Workshop 'User Interface for All'*. Retrieved June 6 2006, from ACM database.

Kobayashi, M., Takeda., K.(2000). *Information Retrieval on the Web*. Retrieved June 5, 2006, from ACM database.

Murugesan, S., Deshpande., Hansen., & Ginige, A. (2000) *Web Engineering: A new Discipline for Development of Web-based Systems*. Retrieved April 11, 2006, from ACM database.

- Nijaz, B. (2000). Web-based Information Access for Agile Management. *International Journal of Agile Management Systems*. 2(2), 121-129.
- Novell Migrating Study (2004): Migrating From IIS Web Server to Apache on SUSE LINUX Enterprise Server 9. Retrieved August 23, 2006, from <http://www.google.com>.
- Replix EDM v3.0 (2005) ReplixEDM: *Web-based Document Management Solution that Enables Enterprises to Manage the Information They Depend On*. Retrieved June 6, 2006, from <http://demox7.softlinx.com/admin>.
- Tang, H., Wu, Y., Yao, J.T., Wang, G., & Yao, Y.Y (2003). *CUPTRSS: A Web-based Research Support System*. Retrieved August 22, 2006, from CiteSeer database.
- The Rose JADE Link. (2003). Jade Software Software Corporation. Retrieved August 26, 2005, from <http://www.jadeworld.com/downloads/jade6/RoseJADELlink.pdf>.
- Uimonen, T. (2000). Case: rational rose. 81940 A Seminar on Reverse Engineering. Retrieved August 26, 2005, from <http://www.cs.tut.fi/~tsysta/sem/Reports/Rose.pdf>.
- Yao, Y.Y (2002). *Information Retrieval Support Systems*. Retrieved August 22, 2006, from CiteSeer database.
- Yao, J.T., & Yao, Y.Y. (2003). *Information Granulation for web-based Information Retrieval Support Systems*. Retrieved August 19, 2006 from <http://www.google.com>

Yao, J.T., & Yao, Y.Y. (2003). *Web-based Information Retrieval Support Systems: building research tools for scientists in the new information age*. Retrieved August 22, 2006 from <http://www.google.com>

Zou, Y., & Kontogiannis, K. (2000). *Web-Based Specification and Integration of Legacy Services*. Retrieved June 6, 2006, from ACM database.

Zhao, W. (1999). *A Study of Web-based Application Architecture and Performance Measurement*. Retrieved June 7, 2006, from Citeseer database.